ADVANCES IN GENERALIZED LINEAR MIXED-EFFECTS MODELS

P. D. Bliese
US Army Medical Research Unit – Europe
Walter Reed Army Institute of Research
APO, AE 09042

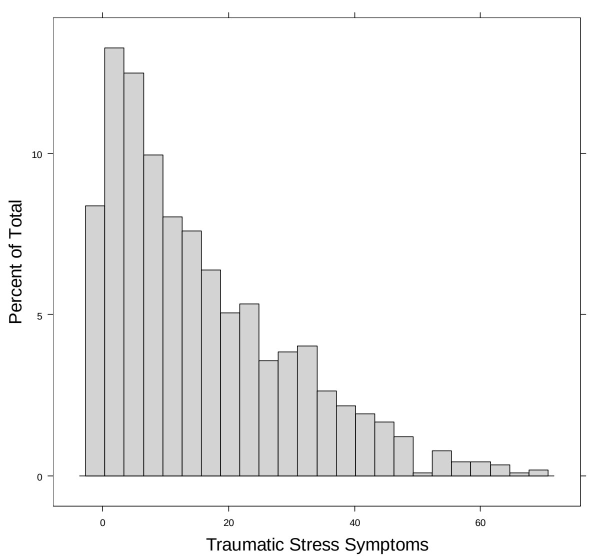
ABSTRACT

Analyses of applied military data is often complex for two reasons. First, many performance and health variables of interest are non-normally distributed. Second, data typically have a complex nested or partially crossed structure. Recent advances in applied statistics make it possible to address both complexities within a single unified statistical framework.

1. INTRODUCTION

Many performance and health variables of interest to commanders, policy makers and researchers are non-normally distributed. For instance, retention, depression and some forms of performance are dichotomous – a Soldier stays or leaves active duty; is or is not depressed; and succeeds or fails on a specific task. Other outcomes have a Poisson distribution. Reports of traumatic stress symptoms, for example, typically have the majority of respondents endorsing few or no symptoms with symptomatic individuals producing a positive skew (see Figure 1). Data with these characteristics obviously violate the assumption of normally distributed data underlying many statistical techniques.

Figure 1: Histogram of Traumatic Stress Symptoms



1.1 Generalized Linear Models

Statistical models to address non-normally distributed data are well-developed. A class of models called Generalized Linear Models or GLMs provide a common framework for analyzing a range of normally and non-normally distributed data. These are described in detail by McCullagh and Nelder (1989) among others. GLMs are implemented in a variety of statistical packages to include the open-source language R (R Development Core Team, 2006).

While somewhat oversimplified, users specify the appropriate family from which the outcome is assumed to have been drawn (binomial, Poisson, normal). In turn, the GLM programs use appropriate corresponding link functions (logit, log, identity) and estimate the model. Model parameters, standard error estimates, and significance values are thereby based on appropriate underlying distribution assumptions.

GLM programs are highly flexible; however, one of their underlying assumptions is the supposition that observations are independent. This assumption leads to an expectation that any residual errors in the model will also be independent. If non-independence is suspected, it is generally included as a fixed (non-random) predictor in the model.

For military analysts, the assumption of independence and/or the need to model non-independence as a fixed-effect potentially represent a fairly substantial limitation. This is because military data often contain non-independence due to either unit membership or repeated measures.

1.2 Mixed-Effects Models

Non-independence due to groups is pervasive in military data because unit (Battalion, Company, Platoon, Squad) membership often influences Soldier-level data (Bliese, 2006). That is, Soldiers' responses are partially a function of the group to which the Soldier belongs. If non-independent data are treated as being independent, one runs the risk of biasing standard errors and reaching

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 NOV 2006**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Advances In Generalized Linear Mixed-Effects Models**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US Army Medical Research Unit Europe Walter Reed Army Institute of Research APO, AE 09042**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002075., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**5**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

incorrect conclusions about what is and is not significant (Bryk & Raudenbush, 1992; Bliese & Hanges, 2004).

In GLM models, the impact of the non-independent data structure can be partially controlled and modeled by including fixed effects for group membership in the form of N-1 dummy codes where N represents the number of groups. Statistically, however, using dummy codes is less than optimal for three reasons. First, the use of dummy codes is "expensive" in terms of degrees of freedom particularly when group sizes are small and there are a lot of groups relative to the total sample size. Second, the use and interpretation of dummy codes is most useful in cases where designs are balanced (i.e., equal group sizes). In practice, though, group sizes often vary considerably either because certain groups are larger than others (HHC company versus Armor company) or because the number of Soldiers representing the unit vary.

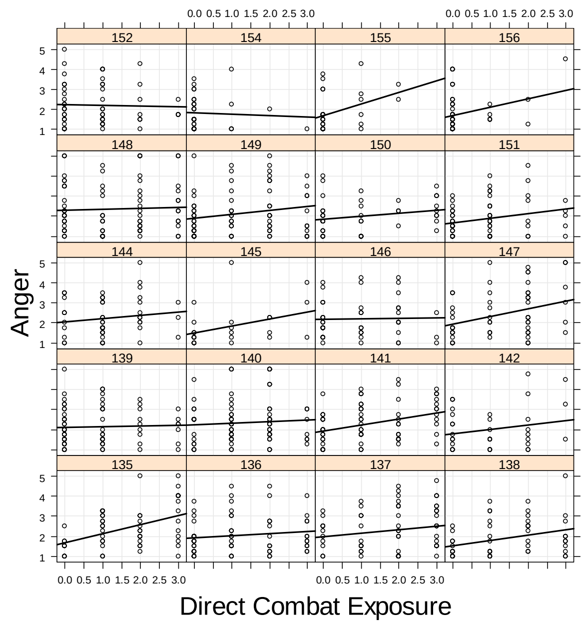
The third reason dummy codes are a less than optimal solution is that on theoretical grounds the specific groups in the sample are often considered to be randomly drawn from a larger population of groups. As such, one is rarely interested in making inferences about any specific group (e.g., Alpha company). Representing groups as N-1 dummy codes, however, treats groups as fixed effects and thereby results in a fixed-effect estimate for each group.

Mixed effects models (also referred to as random effects models) provide a way to control for non-independence. The approach uses few degrees of freedom; easily handles non-balanced designs (i.e., unequal group sizes), and is congruent with the idea that the groups in the sample reflect a random draw from a larger population of groups (Bryk & Raudenbush, 1992; Pinheiro & Bates, 2000).

In addition to efficiently handling non-independence, mixed-effects models also provide a way to extend the substantive understanding of the models by including (a) predictors of group differences and (b) predictors that explain why individual-level relationships vary across groups.

With respect to variability in individual-level relationships, Figure 2, for instance, shows that the relationship between Soldier reports of direct combat exposure and reports of anger-related behaviors is positive (anger increases as reports of combat exposure increase). Across the 20 groups plotted in the figure, however, the strength of the relationship appears to vary. Mixed effect models provide a way to formally test whether the slope variation is reliable or whether it represents random variability. If the slope differences are reliable, the models provide a way to add group-level predictors as explanatory variables. Group-level variables in this situation often refer to some aspect of social context (Bliese, 2006).

Figure 2: Combat and Anger by Unit



Mixed-effects models are also useful in cases where individuals provide repeated measures over time. In longitudinal studies the individual is analogous to the group, and the repeated measures are analogous to individual observations within the group. The inclusion of time adds some additional complexity (Bliese & Ployhart, 2002; Pinhiero & Bates, 2000); however, the heuristic of considering repeated measures from individuals to be equivalent to measures from multiple individuals in groups is helpful nonetheless.

The major limitation with mixed-effects models is that they have been primarily developed for the analysis of normally distributed outcome data (Pinheiro & Bates, 2000). Thus, despite their appeal, they have been somewhat limited in terms of their applicability in applied military settings where outcomes are in many cases not normally distributed.

2. GLMM MODELS

In 2003, the US Army Medical and Material Command (USAMRMC) used the Small Business Technology Transfer Program (STTR) program to explore the possibility of developing new computation methods to facilitate merging the areas of GLM and Mixed-Effects Models – an area referred to as Generalized Linear Mixed effects Models or GLMMs.

A phase II STTR award was granted to TOYON Research Corporation and the University of Wisconsin.

Dr. Douglas Bates from the Department of Statistics at the University of Wisconsin led the development of computational methods.

The original focus of the STTR centered on improving algorithms for estimating statistical parameters in cases where GLMMs were applied to dichotomous data. In 2003, algorithms to analyze dichotomous data within a mixed-effects framework existed, but there were questions about the accuracy of the parameter estimates and standard errors from these algorithms (Rodriguez & Goldman, 2001). Thus, the initial focus of the STTR was to investigate the feasibility of incorporating new algorithms to improve the accuracy of model estimation.

Two promising options included algorithms based on Laplace and Adaptive Quadrature estimation. The initial work focused on how these computationally demanding algorithms could practically be incorporated onto an existing software platform for PCs. That is, while these alternative methods seemed mathematically feasible, the practical challenge of implementing them on a PC-based software platform to test their performance was challenging.

In 2004, while working on ways to implement efficient algorithms into the open source platform R, the researchers observed that the challenges associated with estimating model parameters in GLMMs shared similarities with sparse matrix estimation problems. Based on this observation, the researchers began to explore whether the potential computational advantages of approaching the estimation as a sparse matrix problem could be applied to GLMMs. The sparse matrix estimation approach turned out to be highly feasible and represented a novel computation solution.

In 2004 and 2005, the STTR team implemented the solution into the open-source statistical platform, R (R Development Core Team, 2006) as part of the lme4 package. By 2006, the lme4 package had received extensive testing from a host of R users with a variety of different analytic problems.

In addition to providing a platform for estimating GLMMs on PCs, the STTR produced at least one unexpected, but highly valuable by-product. Namely, the group working on the STTR discovered that when GLMM models were estimated using sparse matrix estimation approaches, the flexibility of the models for handling nested, fully crossed and partially crossed designs greatly increased.

This finding was important because mixed-effects models are generally optimized to handle completely nested designs. For instance, the original mixed-effects models in R (Pinheiro and Bates, 2000) are designed to handle multiple levels of nesting (squad nested in platoon nested in company nested in battalion). In this fully nested design, all members of a single lower-level group are also members of the same higher-level hierarchical group. For instance, all platoon members in a specific platoon are also members of the same company and the same battalion.

In fully nested designs, the mixed-effects models provide separate variance estimates for each level. In this way, one might observe, for example, that most of the group-level variance surrounding reports of anger can be attributed to the platoon level. The ability to parse variance into separate hierarchical levels is useful for understanding what types of variables relate to a various outcomes – finding large variance differences at the platoon-level suggests platoon-level variables impact the specific outcome.

In practice, however, military data is not always cleanly nested. The lack of full nesting occurs when subordinate units are task organized under other units. Task organization leads to situations where two different platoons in the same company may consider themselves to be from two different battalions.

Another situation where partial crossing is common is in longitudinal studies where one is modeling repeated measures nested within individuals and individuals are nested within groups, but during the course of the study the individuals un-uniformly change groups. This can easily happen in military research, and is also common in educational research where students change classrooms and schools over time, but where different students go to different classrooms and different schools.

For traditional mixed-effects models, these partially crossed situations make it virtually impossible to obtain unique variance estimates for each level of the design. If the models are estimated using the sparse matrix approach underlying lme4, however, partially crossed designs pose no problem.

3. PRACTICAL EXAMPLES

While the lme4 package still continues to be refined, its use can be illustrated in several examples.

3.1 Variance Decomposition

For a practical example of the utility of the lme4 approach in variance decomposition, consider a recent randomized controlled trial (RCT) conducted at the US Army Medical Research Unit – Europe (USAMRU-E). In the RCT, platoons were randomly assigned to one of three conditions, and the conditions were indexed by session number.

In an ideal situation, platoon membership should have been completely overlapping with session number. That is, individuals would have been members of one and only one platoon; every platoon member would have been in the same session, and every member of a platoon and session would have been a member of the same company.

In practice, however, individual platoon members were occasionally split so that some members ended up in one session while other members ended up in other sessions. In addition, because of actual work characteristics and the desire to keep intact workgroups together, sessions had the potential to be composed of individuals who were technically in different platoons (e.g., HHC units).

Because pre-existing group differences could bias the results of the RCT, it was necessary to control for these group differences. It was not clear, however, whether the lowest level should have been controlled by indexing session or platoon. In traditional mixed-effects models, the partial crossing would have made a variance decomposition of the separate effects for session versus platoon impossible. In lme4, however, it was possible to estimate variance effects for platoon, session and company and use the results of the variance estimates to control for the grouping which accounted for the most variance.

3.2 Combat and Adjustment

A second example of where the lme4 algorithms are being applied is in large studies of US Army Soldiers returning from Iraq. One of the questions in these data centers on the degree to which reports of specific combat experiences relate to reports of adjustment problems (traumatic stress, depression, alcohol use, etc.).

Many of the adjustment issues such as traumatic stress and depression can be measured both in terms of symptom severity (usually a Poisson distribution) and whether or not a score meets or exceeds cut-off criteria (binomial).

The GLMM routines in lme4 provide a way to determine whether social context in the military plays a role in how reports of specific combat experiences relate to adjustment problems. For instance, if there is unit-by-unit variability in the strength of the relationship between seeing dead or dying combatants and showing symptoms of traumatic stress, it suggests that unit level factors may ameliorate the effects of this combat experience. This, in turn, may provide suggestions for early interventions. A detailed discussion of the role of unit-level factors in non-combat situations can be found in Bliese and Jex (2002), but this work suggests the importance of the role social context in combat situations.

3.3 Sleep and Performance

Studies of individual performance under conditions of sleep restriction and sleep deprivation reveal consistent and reliable individual differences in vulnerability to the effects of sleep deprivation (Bliese, Wesensten, & Balkin, in press; Van Dongen, Maislin, Mullington, & Dinges, 2003). One of the potentially fruitful areas for future research is to identify individual factors that predict vulnerability to sleep loss.

Both Bliese et al. (in press) and Van Dongen et al. (2003) used mixed effects models to investigate individual differences. In both cases the psychomotor vigilance task was used in longitudinal experiments; however, in both cases the modeling of the variable relied on outcome summaries of task performance. Specifically, Bliese et al. (in press) used average reaction time across day excluding lapses, while Van Dongen et al. (2003) analyzed a sum of lapses averaged across days. In both cases, this was done to create an outcome with normal distribution properties.

Advances in GLMMs make it possible to model specific trial performance on binomial tasks and other tasks which may not provide normally distributed outcomes. This, in turn, may facilitate the ability to examine trial-by-trail performance.

4. CONCLUSIONS

Innovations in GLMMs and the availability of the lme4 program in the open-source language R (R Development Core Team, 2006) open the possibility of analyzing a wide-range of military data for informing policy decisions and for advancing scientific knowledge. The merge of GLM models and mixed-effects models provide a set of statistical techniques that are specifically designed to address the complex nature of military data.

ACKNOWLEDGEMENTS

The views expressed here are those of the authors and do not reflect the view of the Department of the Army, Department of Defense, or the US government.

REFERENCES

Bliese, P.D., 2006: Social climates: Drivers of soldier well-being and resilience. In A.B. Adler, C.A. Castro, & T.W. Britt (Eds.), *Military life: The psychology of serving in peace and combat*: Vol. 2.

Operational Stress (pp. 213-234). Westport, CT: Praeger Security International.

Bliese, P. D. & Hanges, P. J. 2004: Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. *Organizational Research Methods*, **7**, 400-417.

Bliese, P. D. & Jex, S. M. 2002: Incorporating a multilevel perspective into occupational stress research: Theoretical, methodological, and practical implications. *Journal of Occupational Health Psychology*, **7**, 265-276.

Bliese, P. D., Wesensten, N. J., & Balkin, T. J., in press: Age and Inter-individual Variability in Performance during Sleep Restriction. *Journal of Sleep Research*.

Bryk, A. P., & Raudenbush, S. W. 1992. *Hierarchical linear models*. Newbury Park, CA: Sage.

Pinheiro, J. C. & Bates, D. M. 2000: *Mixed-effects models in S and S-PLUS*. New York: Springer-Verlag.

McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models* (2^{nd} Ed). London: Chapman & Hall.

R Development Core Team, 2006: *R: A language and environment for statistical computing, reference index version 2.3.1*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rodriguez, G. & Goldman, N. 2001: Improved estimation procedures for multilevel models with binary response: A case study. *Journal of the Royal Statistical Society*, **164**, 339-355.

Van Dongen, H. P. A., Maislin, G., Mullington, J. M., Dinges, D. F., 2003: The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep*, **26**: 117-126